

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-266385

(43)Date of publication of application : 22.09.1994

(51)Int.Cl.

G10L 3/00

G10L 3/00

(21)Application number : 05-055541

(71)Applicant : FUJITSU LTD

(22)Date of filing : 16.03.1993

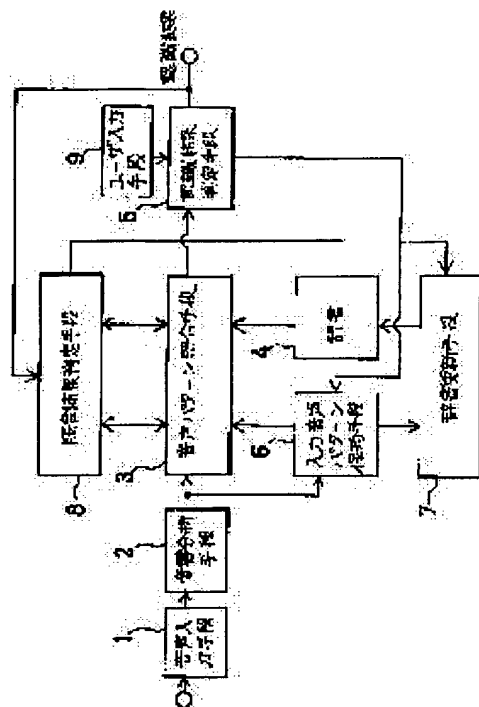
(72)Inventor : HAMAZAKI RYOSUKE

(54) DICTIONARY UPDATE OF SPEECH RECOGNITION DEVICE

(57)Abstract:

PURPOSE: To maintain a high recognition rate even if the vocalization state of a speaker changes and to evade the registration of a distorted speech pattern in a dictionary.

CONSTITUTION: A speech pattern collating means 3 finds the score between an input speech pattern whose featured are extracted by an acoustic analyzing means 2 and the template in the dictionary 4 and the score to a last correct answer input speech pattern held in an input speech pattern holding means 6. A recognition result decision means 5 outputs a recognition result on the basis of the scores. Further, when the recognition result is not correct, a user input means 9 grants a correct answer label to the input speech pattern. A collating result decision means 8 compares the score between the dictionary 4 and correct answer input speech pattern with the score between the correct answer speech pattern in the input speech pattern holding means 6 and the input speech pattern. A dictionary update means 7 updates the dictionary 4 on the basis of the comparison result.



LEGAL STATUS

[Date of request for examination]	22.02.2000
[Date of sending the examiner's decision of rejection]	17.09.2002
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	3514481
[Date of registration]	23.01.2004
[Number of appeal against examiner's decision of rejection]	2002-20187
[Date of requesting appeal against examiner's decision of rejection]	17.10.2002
[Date of extinction of right]	

Copyright (C); 1998,2003 Japan Patent Office

BEST AVAILABLE COPY

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] voice input means (1) Sonagraphy means (2) which carries out [voice / which was inputted / strange input] sonagraphy sonagraphy means (2) the obtained input voice pattern -- beforehand -- dictionary (4) A voice pattern-matching means (3) to collate the standard voice pattern corresponding to each label registered into inside a recognition result judging means (5) to obtain a recognition result based on the collating result A user input means to give the label of a correct answer to an input voice pattern (9) It is a dictionary (4) by the input voice pattern. A renewal means of a dictionary to update (7) In the voice recognition unit which it had an input voice pattern maintenance means (6) to hold an input voice pattern temporarily the input voice pattern whose recognition result it prepared and was a correct answer -- or It is a user input means (9) at the time of recognition. About the input voice pattern with which the correct answer label was given, it is a voice pattern-matching means (3). It sets. An input voice pattern and dictionary (4) While collating each registered correct answer standard voice pattern Input voice pattern maintenance means (6) When the correct answer voice pattern to which the same label as the above-mentioned input voice pattern was given is held The above-mentioned input voice pattern and input voice pattern maintenance means (6) The held voice pattern is collated. The above-mentioned input voice pattern and input voice pattern maintenance means (6) The similarity of the held voice pattern The above-mentioned input voice pattern and dictionary (4) When larger than similarity with the registered standard voice pattern Input voice pattern maintenance means (6) It is a dictionary (4) about the held voice pattern. It registers and is a dictionary (4). Renewal method of a dictionary in the voice recognition unit characterized by deleting a standard voice pattern.

[Claim 2] voice input means (1) Sonagraphy means (2) which carries out [voice / which was inputted / strange input] sonagraphy sonagraphy means (2) the obtained input voice pattern -- beforehand -- dictionary (4) A voice pattern-matching means (3) to collate the standard voice pattern corresponding to each label registered into inside a recognition result judging means (5) to obtain a recognition result based on the collating result A user input means to give the label of a correct answer to an input voice pattern (9) It is a dictionary (4) by the input voice pattern. A renewal means of a dictionary to update (7) In the voice recognition unit which it had an input voice pattern maintenance means (6) to hold an input voice pattern temporarily preparing -- voice pattern-matching means (3) the input voice pattern whose recognition result it set and was a correct answer -- or About the input voice pattern by which the correct answer label was given with the user input means at the time of recognition, they are an input voice pattern and a dictionary (4). While collating each registered correct answer standard voice pattern Input voice pattern maintenance means (6) When the correct answer voice pattern to which the same label as the above-mentioned input voice pattern was given is held The above-mentioned input voice pattern and input voice pattern maintenance means (6) The held voice pattern is collated. The above-mentioned input voice pattern and input voice pattern maintenance means (6) The similarity of the held voice pattern The above-mentioned input voice pattern and dictionary (4) The input voice pattern newly inputted when larger than similarity with the registered standard voice pattern is registered into a dictionary (4), and it is

a dictionary (4). Renewal method of a dictionary in the voice recognition unit characterized by deleting a standard voice pattern.

[Claim 3] A new input voice pattern and input voice pattern maintenance means (6) The similarity of the held voice pattern The above-mentioned input voice pattern and dictionary (4) When smaller than similarity with the registered standard voice pattern Input voice pattern maintenance means (6) The held voice pattern is deleted and it is an input voice pattern maintenance means (6) about an input voice pattern. Renewal method of a dictionary in the voice recognition unit of claim 1 characterized by registering, or claim 2.

[Claim 4] A new input voice pattern and input voice pattern maintenance means (6) The similarity of the held voice pattern The above-mentioned input voice pattern and dictionary (4) When smaller than similarity with the registered standard voice pattern Input voice pattern maintenance means (6) Renewal method of a dictionary in the voice recognition unit of claim 1 characterized by leaving the held voice pattern as it is, and deleting the above-mentioned input voice pattern, or claim 2.

[Claim 5] A new input voice pattern and an input voice pattern maintenance means (6) The similarity of the held voice pattern is the above-mentioned input voice pattern and a dictionary (4). When small than similarity with the registered standard voice pattern, it is an input voice pattern maintenance means (6). The renewal method of a dictionary in the voice recognition unit of claim 1 characterized by to delete the held voice pattern and an input voice pattern, or claim 2.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to the renewal method of a dictionary in the voice recognition unit in which especially this invention had the means which updates a dictionary by the input voice pattern (relearning) about the renewal method of a dictionary in a voice recognition unit.

[0002]

[Description of the Prior Art] In the voice recognition unit of a registration mold, before beginning to use recognition equipment, those who use recognition equipment beforehand need to register the description of their voice into the dictionary of a voice recognition unit. However, the description of voice tends to change, and if it changes with the passage of time and it not only may change easily, but is continuing using the same dictionary by the environment of the perimeter at the time of recognition, or physical and mental limit, a recognition rate will fall gradually.

[0003] In order to cope with this conventionally, the dictionary was updated by replacing a correct answer template with furthest input configuration and distance by which sonagraphy was carried out at the time of recognition when it was a correct answer to the dictionary of a multi-template. Drawing 7 is drawing showing the above-mentioned conventional voice recognition unit. A voice input means by which 1 inputs voice with a microphone etc. in this drawing, A sonagraphy means for 2 to perform frequency analysis etc. to the sound signal inputted by the microphone etc., and to extract the description as voice, and the input voice pattern as which 3 was analyzed by the sonagraphy means 2, The "similarity" which shows how many two patterns collating with each template which is analyzed beforehand and registered into the dictionary is performed, and are alike, or the "distance" which shows whether it is how much separated is calculated (these are hereafter called score), and it is a voice pattern-matching means to output.

[0004] Moreover, the dictionary in which 4 registered the pattern with which the candidate for recognition was analyzed beforehand, and 5 follow the score to each template outputted from the voice pattern-matching means 3. A recognition result judging means by which sort and similarity outputs one or more template labels to descending or order with a small distance, An input voice pattern maintenance means to hold temporarily the input voice pattern as which 6 was analyzed by the sonagraphy means 2, and 7 are renewal means of a dictionary to register an input voice pattern into a dictionary 4, or to delete the template of a dictionary 4.

[0005] In this drawing, it is given to the input voice pattern maintenance means 6, and the voice inputted from the voice input means 1 is held while being analyzed by the sonagraphy means 2, extracting the description and being given to the voice pattern-matching means 3. The voice pattern-matching means 3 collates the template registered into the dictionary 4, and the feature parameter of input voice, and asks for the score between them. The recognition result judging means 5 judges input voice based on the score which the input and the voice pattern-matching means 3 from a user output, and outputs a recognition result.

[0006] Moreover, the output of the recognition result judging means 5 is given to the renewal means 7 of a dictionary, and the renewal means 7 of a dictionary updates a dictionary by replacing a correct answer template with the furthest distance registered into the dictionary 4, and the input voice pattern held at the input voice pattern maintenance means 6, when a recognition result is a correct answer.

[0007]

[Problem(s) to be Solved by the Invention] By the way, in the above-mentioned conventional renewal method of a dictionary, when a recognition result is a correct answer, the input voice pattern held at the input voice pattern maintenance means 6 and a correct answer template with the furthest distance of a forward recognition category are replaced, without performing a check in any way.

[0008] For this reason, even if it was the case where un-arranging -- an input voice pattern is distorted by chance -- had arisen, in forward recognition, there was a problem that a dictionary will be updated. Moreover, although it may be the natural result of being based on change of vocal quality even when it has incorrect-recognized, in the above-mentioned conventional renewal method of a dictionary, in such a case, renewal of a dictionary was not performed, but there was a problem that it was continued by restraining an initial template.

[0009] Drawing 8 and drawing 9 are the conceptual diagrams of the description parameter space at the time of changing vocal quality, and drawing 8 and drawing 9 show the case where three categories A, B, and C exist. In drawing 8 and drawing 9, drawing 8 shows the case where distribution of the category of dictionary creation time and distribution according [drawing 9 (a) and (b)] to fluctuation of vocal quality are changed. O The place [surrounding] is the range over which each category is distributed, in drawing 8, the sample of a black trigonum, a black dot, and a black rectangular head exists in it, and the input sample at the time of recognition is shown by the black rectangular head and the black dot in drawing 9 (a) and (b). Moreover, the boundary of three categories is shown by the continuous line and the dotted line shows signs that the boundary line moved by change of vocal quality.

[0010] As shown in drawing 8, after drawing up a dictionary, since the distribution condition of each category at the time of recognition and it of a dictionary are in agreement, they can acquire a high recognition rate for the time being. Here, it is assumed with the passage of time that it is that from which the distribution condition of each actual category changed as shown in drawing 9 (a) and (b) by change of vocal quality etc. In addition, although it assumes that distribution of the expedient top of explanation and a category moved in the direction same on the whole, it is changing to this drawing actual more intricately.

[0011] When distribution of each category moves, the boundary line of each actual category moves from the continuous line to the dotted line, as shown in drawing 9 (a) and drawing 9 (b). Now, in drawing 9 (a), a certain sample shown at a black rectangular head is observed. This input sample presupposes that it is category C' in fact. This input sample is widely different from distribution of category C' after time amount progress, and should be essentially judged by the boundary of a dotted line to be Category B. However, in distribution of the dictionary creation time shown in drawing 8, it will be judged with Category C and a dictionary will be updated by the above-mentioned sample in the above mentioned conventional method.

[0012] Moreover, in drawing 9 (b), a certain sample shown by the black dot is observed. This input sample presupposes that it is category B' in fact. Although it should be updated since this input sample is originally judged by the boundary of a dotted line to be Category B, in distribution of dictionary creation time, it is judged with Category A and a dictionary is not updated. As mentioned above, in the conventional method, the case where a dictionary is updated although it should not update essentially, and it is not updated although a dictionary should be updated arises.

[0013] Even if this invention is made in order to improve the trouble of the above-mentioned conventional technique, and a speaker's utterance condition changes with the passage of time By updating the description pattern of a dictionary to a new thing, whenever it recognizes It aims at offering the renewal method of a dictionary in the voice recognition unit which can prevent that the voice pattern distorted by chance possible [maintaining a high recognition rate] even if it changes a speaker's utterance

condition is registered into a dictionary.

[0014]

[Means for Solving the Problem] Drawing 1 is the principle explanatory view of this invention. A voice input means by which 1 inputs voice with a microphone etc. in this drawing, A sonagraphy means for 2 to perform frequency analysis etc. to the sound signal inputted by the microphone etc., and to extract the description as voice, and the input voice pattern as which 3 was analyzed by the sonagraphy means 2, While performing collating with each template which is analyzed beforehand and registered into the dictionary Collating with the voice pattern inputted just before being held at the input voice pattern maintenance means 6 is performed. A voice pattern-matching means to calculate similarity or distance, the dictionary, in which 4 registered the pattern with which the candidate for recognition was analyzed beforehand, A recognition result judging means by which 5 outputs a recognition result based on the output of the voice pattern-matching means 3, An input voice pattern maintenance means to hold temporarily the input voice pattern as which 6 was analyzed by the sonagraphy means 2, A renewal means of a dictionary for 7 to register an input voice pattern into a dictionary 4, or to delete the template of a dictionary 4, About the template set as the recognition object of input voice, 8 compares the score by the side of the dictionary 4 outputted from the voice pattern-matching means 3, and the input voice pattern maintenance means 6. A collating result judging means to judge whether the score of what one is expensive, and 9 are user input means to give a correct answer label to an input voice pattern.

[0015] In order to solve the above-mentioned technical problem, invention of claim 1 of this invention The sonagraphy means 2 which carries out [voice / which was inputted from the voice input means 1 / strange input] sonagraphy as shown in drawing 1 , A voice pattern-matching means 3 to collate the input voice pattern obtained by the sonagraphy means 2, and the standard voice pattern corresponding to each label beforehand registered into the dictionary 4, In the voice recognition unit equipped with a recognition result judging means 5 to obtain a recognition result, a renewal means 7 of a dictionary to update a dictionary 4 with an input voice pattern, and a user input means 9 to give the label of a correct answer to an input voice pattern, based on the collating result the input voice pattern whose recognition result an input voice pattern maintenance means 6 to hold an input voice pattern temporarily was established, and was a correct answer -- or While collating each correct answer standard voice pattern registered into the input voice pattern and the dictionary 4 in the voice pattern-matching means 3 about the input voice pattern by which the correct answer label was given with the user input means 9 at the time of recognition When the correct answer voice pattern with which the same label as the above-mentioned input voice pattern was given to the input voice pattern maintenance means 6 is held The above-mentioned input voice pattern and the voice pattern held at the input voice pattern maintenance means 6 are collated. The similarity of the above-mentioned input voice pattern and the voice pattern held at the input voice pattern maintenance means 6 The voice pattern which came size and was held for the input voice pattern maintenance means 6 at the case from the similarity of the above-mentioned input voice pattern and the standard voice pattern registered into the dictionary 4 is registered into a dictionary 4, and the standard voice pattern of a dictionary 4 is deleted.

[0016] The similarity of a new input voice pattern and the voice pattern held at the input voice pattern maintenance means 6 invention of claim 2 of this invention When larger than the similarity of the above-mentioned input voice pattern and the standard voice pattern registered into the dictionary 4 The newly inputted input voice pattern is registered into a dictionary 4 instead of registering into a dictionary 4 the voice pattern held like invention of invention of claim 1 at the input voice pattern maintenance means 6.

[0017] Invention of claim 3 of this invention is set to invention of claim 1 or claim 2. The similarity of a new input voice pattern and the voice pattern held at the input voice pattern maintenance means 6 When smaller than the similarity of the above-mentioned input voice pattern and the standard voice pattern registered into the dictionary 4, the voice pattern held at the input voice pattern maintenance means 6 is deleted, and an input voice pattern is registered into the input voice pattern maintenance means 6.

[0018] When the similarity of a new input voice pattern and the voice pattern held at the input voice

pattern maintenance means 6 is small than the similarity of the above-mentioned input voice pattern and the standard voice pattern registered into the dictionary 4, invention of claim 4 of this invention leaves the voice pattern held at an input voice pattern maintenance means 6 as it is, and deletes the above-mentioned input voice pattern in invention of claim 1 or claim 2.

[0019] Invention of claim 5 of this invention deletes the voice pattern and the input voice pattern held at the input voice pattern maintenance means 6 in invention of claim 1 or claim 2, when the similarity of a new input voice pattern and the voice pattern held at the input voice pattern maintenance means 6 is small than the similarity of the above-mentioned input voice pattern and the standard voice pattern registered into the dictionary 4.

[0020]

[Function] In drawing 1, it is analyzed by the sonagraphy means 2, the description is extracted, and the voice inputted from the voice input means 1 is given to the voice pattern-matching means 3. While the voice pattern-matching means 3 performs collating with each standard voice pattern which is beforehand analyzed as the input voice pattern analyzed by the sonagraphy means 2, and is registered into the dictionary 4. When the recognition result of an input voice pattern is a correct answer, or when a correct answer label is given to an input voice pattern by the user input, collating with the correct answer voice pattern inputted just before being held at the input voice pattern maintenance means 6 is performed, and a score is calculated.

[0021] The recognition result judging means 5 judges input voice based on the score which the voice pattern-matching means 3 outputs, and outputs a recognition result. Moreover, when a recognition result is an unjust solution, a correct answer label is given to an input voice pattern based on the user input from the user input means 9. The collating result of the correct answer standard voice pattern of a dictionary 4 and an input voice pattern and the collating result of the voice pattern and the input voice pattern with which the correct answer label currently held at an input voice pattern maintenance means 6 was given compare about the input voice pattern whose recognition result was a correct answer, or the input voice pattern by which a correct answer label was given with a user input means 9 at the time of recognition, and a collating result judging means 8 judges [that it is high in which score, and].

[0022] As a result of judging with the collating result judging means 8, the direction of the collating result of the voice pattern and input voice pattern which are held at the input voice pattern maintenance means 6 the renewal means 7 of a dictionary. When higher than the score of the collating result of a dictionary 4 and an input voice pattern, a minimum score standard voice pattern is deleted from a dictionary 4. A standard voice pattern is replaced by registering into a dictionary 4 the voice pattern currently held at the input voice pattern maintenance means 6, or an input voice pattern.

[0023] As mentioned above, it sets to invention of claim 1 of this invention thru/or claim 5. When an input voice pattern maintenance means 6 to hold the input voice which functions as a temporary dictionary apart from the dictionary currently prepared beforehand is established and the same voice pattern as the 2nd times is inputted at the time of renewal of the dictionary 4 by the input voice pattern the input voice pattern whose recognition result was a correct answer -- or The input voice pattern to which the correct answer label was given by the user input means 9, It collates about each with the correct answer standard voice pattern registered into the dictionary 4, and the voice pattern with which the correct answer label currently held temporarily was given to the input voice pattern maintenance means 6. The dictionary 4 is updated when the score between an input voice pattern and the voice pattern held at the input voice pattern maintenance means 6 is more expensive among those collating results.

[0024] Therefore, rather than the correct answer standard voice pattern of the dictionary 4 in which the past was created, naturally it is expected that a score will become [the direction of the correct answer voice pattern inputted just before being held at the input voice pattern maintenance means 6] high, and it can update a dictionary 4 corresponding to fluctuation of a speaker's utterance condition. The voice pattern currently held here at the input voice pattern maintenance means 6 inputted immediately before. Although the renewal of a dictionary 4 must be avoided when it becomes what was distorted by the background noise or the fault of utterance, it sets to invention of claim 1 of this invention thru/or claim 5.

The correct answer standard voice pattern registered into the dictionary 4 in the correct answer input voice pattern as mentioned above, It collates with the input voice pattern maintenance means 6 about each with the correct answer voice pattern currently held temporarily. Since the dictionary 4 is updated when the score between an input voice pattern and the voice pattern held at the input voice pattern maintenance means 6 is more expensive among those collating results As long as there is no fault in the voice uttered by the 2nd times, it is judged with the standard voice pattern of a score registered into the dictionary 4 being more expensive than the bent voice pattern held at the input voice pattern maintenance means 6, and renewal of the dictionary 4 by the bent voice pattern can be avoided.

[0025] Furthermore, it sets to invention of claim 2 of this invention. Since the newly inputted correct answer input voice pattern was registered into the dictionary 4 instead of registering into a dictionary 4 the correct answer voice pattern held at the input voice pattern maintenance means 6 While being able to acquire the same effectiveness as invention of claim 1, the voice pattern uttered next can be held as it is for the input voice pattern maintenance means 6, and a processing man day can be decreased.

[0026] Furthermore, it sets to claim 3 of this invention thru/or invention of 5. The score between a new correct answer input voice pattern and the correct answer voice pattern held at the input voice pattern maintenance means 6 When smaller than the score between the above-mentioned input voice pattern and the standard voice pattern registered into the dictionary 4 Since the voice pattern held at the input voice pattern maintenance means 6, an input voice pattern, or both were deleted, the bent voice pattern is not held at the input voice pattern maintenance means 6.

[0027]

[Example] The A-D converter which drawing 2 is drawing showing the 1st example of this invention, and changes into a digital signal the voice into which 11 is inputted from a microphone etc. in this drawing, A sonagraphy means for 21 to analyze the sound signal changed into the digital signal by A-D converter 11, and to extract a feature-parameter time series vector, and the feature-parameter time series vector as which 3 was analyzed by the sonagraphy means 2, It is a voice pattern-matching means to perform collating with each template which is analyzed beforehand and registered into the dictionary, and the voice pattern-matching means 3 collates also with this voice pattern, when there is a voice pattern currently held at the 2nd input voice pattern buffer 62.

[0028] Moreover, they are a recognition result judging means output a recognition result based on the dictionary in which 4 registered the pattern with which the candidate for recognition was analyzed beforehand, and the score to each template to which 5 was outputted from a voice pattern-matching means 3, a score sort means sort the score with which a voice pattern-matching means 3 outputs 51, and a recognition result selection means 52 chooses a right recognition result according to the input from a user, and choose a final recognition result.

[0029] The 1st input voice pattern maintenance means which holds temporarily the correct answer input voice pattern as which 61 was analyzed by the sonagraphy means 2, 62 is 2nd input voice pattern maintenance means which holds temporarily the correct answer input voice pattern held at the 1st input voice pattern maintenance means. The 1st and 2nd input voice pattern maintenance means It has the buffer of the number corresponding to the number of the labels of input voice at least, respectively. The voice label given to the 1st and 2nd input voice pattern maintenance means 61 and 62 from the audio label or audio user input at the time of the correct answer given from a recognition result judging means is given, and an input voice pattern is held.

[0030] A renewal means of a dictionary for 7 to register an input voice pattern into a dictionary 4, or to delete the template of a dictionary 4, A voice pattern registration means to register into a dictionary 4 the voice pattern with which 71 was held at the 2nd input voice pattern maintenance means 62, A voice pattern deletion means by which, as for 72, a score deletes the minimum voice pattern from a dictionary 4, 8 compares the score by the side of the dictionary 4 outputted from the voice pattern-matching means 3, and the 2nd input voice pattern maintenance means 62 about the correct answer template of input voice, and is a collating result judging means to judge whether the score of what one is expensive.

[0031] Next, actuation of the 1st example of this invention shown in drawing 2 is explained. Voice is

inputted into the AD translation section 11 from a microphone etc., is changed into a digital signal and sent to the sonagraphy means 21 as discretized signal data. The sonagraphy means 21 extracts an audio feature-parameter time series vector from the signal data by which discretization was carried out [above-mentioned], for example for every fixed time amount of 5msec(s) - 25msec.

[0032] The following technique etc. is known as the extract technique of an audio feature-parameter time series vector.

** What extracts the spectrum in a different frequency band by two or more filter banks.

** What searches for the SU ** KUTORU power time series divided into two or more channels after performing FFT (high-speed Fourier transformation).

** What performs line type prediction analysis (LPC) and searches for the multiplier time series.

** What searches for cepstrum (cepstrum) multiplier time series using FFT (high-speed Fourier transformation) or line type prediction analysis (LPC).

[0033] The feature-parameter time series vector extracted by the sonagraphy means 2 is outputted to the voice pattern-matching means 3, and is collated with the template beforehand registered into the dictionary 4. The DP matching generally used can be used as the technique of collating in the voice pattern-matching means 3, and the voice pattern-matching means 3 asks for the score between 2 patterns as a collating result, and outputs it to the recognition result judging means 5.

[0034] The score sort means 51 of the recognition result judging means 5 performs sorting based on the score calculated with the voice pattern-matching means 3, and a score sorts it in high order. The recognition result selection means 52 chooses a right result according to a user's input, and outputs the final result. On the other hand, the voice pattern whose recognition result was a correct answer among the voice patterns analyzed with the sonagraphy means 2, or the voice pattern with which the correct answer label was given by the user input is outputted also to the 1st input voice pattern buffer 61 while it is outputted to the voice pattern-matching means 3. And if voice new next is inputted, the correct answer voice pattern which the correct answer voice pattern analyzed with the sonagraphy means 2 was held by the 1st input voice pattern buffer 61, and was held at the 1st input voice pattern buffer 61 will be outputted to the 2nd input voice pattern buffer 62, and will be held there. That is, the 1st input voice pattern buffer 61 in this example functions as a buffer for an input voice pattern, and the 2nd input voice pattern buffer 62 functions as a primary dictionary.

[0035] Moreover, in case a voice pattern is held at the 1st and 2nd input voice pattern buffers 61 and 62, as described above, when a correct answer label or a recognition result in case the judgment result by the recognition result judging means 5 is a correct answer is an unjust solution, the correct answer label which a user inputs is given. And a voice pattern-matching means 3 collates the voice pattern with which the correct answer label inputted just before being held at the input voice pattern buffer 62 which functions as the input voice pattern with which the correct answer label was given as a dictionary was given, and asks for a score while it collates the input voice pattern with which the correct answer label was given, and the correct answer template beforehand registered into the dictionary 4.

[0036] The collating result of the correct answer template of a dictionary 4 and an input voice pattern and the collating result of the voice pattern and the input voice pattern with which the correct answer label currently held at an input voice pattern maintenance means 6 was given compare about the input voice pattern whose recognition result was a correct answer, or the input voice pattern by which a correct answer label was given with a user input means 9 at the time of recognition, and a collating result judging means 8 judges [that it is high in which score, and].

[0037] As a result of judging with the collating result judging means 8, the direction of the collating result of the correct answer voice pattern and correct answer input voice pattern which are held at the input voice pattern buffer 62 the renewal means 7 of a dictionary When higher than the score of the collating result of a dictionary 4 and a correct answer input voice pattern, the minimum template of a score is deleted from a dictionary 4, and a template is replaced by registering into a dictionary 4 the correct answer voice pattern currently held at the input voice pattern buffer 62. That is, the template of a dictionary 4 is deleted with the voice pattern deletion means 72 of the renewal means 7 of a dictionary,

and the correct answer voice pattern currently held by the voice pattern registration means 71 at the input voice pattern buffer 62 is registered into a dictionary.

[0038] The collating result of the input voice pattern with which the correct answer label was given in this example at the time of renewal of a dictionary with input voice as explained above, and the correct answer template registered into the dictionary, A collating result with the correct answer voice pattern inputted just before being held at the 2nd input voice pattern maintenance means is compared. Since a dictionary is updated with the voice pattern inputted just before being held at the 2nd input voice pattern maintenance means according to the comparison result, fluctuation of a speaker's utterance condition can be coped with and a dictionary can be updated.

[0039] Moreover, the voice pattern distorted by chance can prevent registering with a dictionary. That is, as long as there is no fault in the voice uttered by the 2nd times, a score becomes high and the template registered into the dictionary 4 rather than the bent voice pattern held at the input voice pattern buffer 62 can avoid renewal of the dictionary 4 by the bent voice pattern.

[0040] Above mentioned drawing 8 and above mentioned drawing 9 explain this point. In drawing 9 (a), the input sample of the above mentioned black rectangular head presupposes once that it was held temporarily at the input voice pattern maintenance means 62 by utterance of an eye. It is avoidable for the direction of the score by the side of a dictionary to become high naturally, and to update a dictionary 4 by the above bent voice patterns by utterance of the voice inputted into the 2nd times being satisfactory here by the collating result of the input voice pattern and the dictionary 4 inputted into the 2nd times when it was in category C', and the collating result with the voice pattern of the above-mentioned sample held at the input voice pattern maintenance means 62.

[0041] Moreover, in drawing 9 (b), the input sample of the above mentioned black dot presupposes once that it was held temporarily at the input voice pattern maintenance means 62 by utterance of an eye. The case where a dictionary 4 can be updated comes out by utterance of the voice inputted into the 2nd times being satisfactory here, and calculating the score of the input voice pattern and the dictionary 4 inputted into the 2nd times although updating was not performed when it was in category B' and having been judged according to the category boundary of dictionary creation time, and a score with the voice pattern of the above-mentioned sample held at the input voice pattern maintenance means 6.

[0042] That is, a dictionary 4 is updated when the voice pattern once inputted into the eye and the score between the furthest samples of Category B are smaller than the score between the voice patterns inputted into the 2nd times as the voice pattern once inputted into the eye. Drawing 3 is drawing showing the 2nd example of this invention. The configuration of this example is fundamentally the same as that of the 1st example shown in drawing 2, and, as for this example and the 1st example, connections between the 1st and 2nd input voice pattern buffers 61 and 62 and the renewal means 7 of a dictionary differ.

[0043] That is, in the 1st example, the voice pattern with which the correct answer label held at the 1st input voice pattern buffer 61 was given is given to the renewal means 7 of a dictionary in the 2nd example to having given the voice pattern with which the correct answer label held at the 2nd input voice pattern buffer 62 was given to the renewal means 7 of a dictionary.

[0044] Therefore, when the collating result of the correct answer voice pattern and the correct answer input voice pattern currently held at the input voice pattern buffer 62 is higher than the score of the collating result of a dictionary 4 and a correct answer input voice pattern, it deletes the minimum template of a score from a dictionary 4, and updates a dictionary 4 in the example of drawing 3 with the voice pattern currently held at the 1st input voice pattern buffer 61.

[0045] In this example, while being able to acquire the same effectiveness as the 1st example, a processing man day can be made fewer than the thing of the 1st example. That is, in this example, since the dictionary 4 is updated with the voice pattern currently held at the 1st input voice pattern buffer 61, by updating a dictionary 4, the contents of the 1st input voice pattern buffer 61 serve as empty, and the correct answer voice pattern uttered next can be held as it is to the 1st input voice pattern buffer 61.

[0046] Drawing 4 is a flow chart which shows the 3rd example of this invention, and the configuration of this example is the same as that of the 1st example shown in drawing 2. This example when the similarity

of a correct answer input voice pattern and the 2nd input voice pattern buffer 62 is larger than the similarity of a correct answer input voice pattern and a dictionary 4. The correct answer description pattern of the 2nd input voice pattern buffer 62 is registered into a dictionary 4. Moreover, when the similarity of a correct answer input voice pattern and the 2nd input voice pattern buffer 62 is smaller, the correct answer description pattern of the 2nd input voice pattern buffer 62 is deleted, and the correct answer description pattern of the 1st input voice pattern buffer 61 is registered into the 2nd input voice pattern buffer 62.

[0047] In this example, since the description pattern currently held at the 2nd input voice pattern buffer 62 has been deleted when the similarity by the side of a dictionary 4 is larger than the similarity by the side of the 2nd input voice pattern buffer 62, when the description pattern of input voice with which input voice was distorted to distortion (deforming) and the 2nd input voice pattern buffer 62 is held, the description pattern can be deleted.

[0048] In addition, in the above-mentioned example, although similarity is used as a score by collating, distance may be used as a score. In that case, the sense of the greater than sign shown in drawing 4 turns into reverse sense. Drawing 5 is a flow chart which shows the 4th example of this invention, and the configuration of this example is the same as that of the 1st example shown in drawing 2.

[0049] This example when the similarity of a correct answer input voice pattern and the 2nd input voice pattern buffer 62 is larger than the similarity of a correct answer input voice pattern and a dictionary 4. The correct answer description pattern of the 2nd input voice pattern buffer 62 is registered into a dictionary 4. Moreover, when the similarity of a correct answer input voice pattern and the 2nd input voice pattern buffer 62 is smaller, the correct answer description pattern of the 1st input voice pattern buffer 61 is deleted.

[0050] In this example, since the description pattern currently held at the 1st input voice pattern buffer 61 has been deleted when the similarity by the side of a dictionary 4 is larger than the similarity by the side of the 2nd input voice pattern buffer 62, when the description pattern of input voice with which input voice was distorted to distortion (deforming) and the 1st input voice pattern buffer 61 is held, the description pattern can be deleted.

[0051] In addition, in the above-mentioned example, although similarity is used as a score by collating, distance may be used as a score like the 3rd example. In that case, the sense of the greater than sign shown in drawing 5 turns into reverse sense. Drawing 6 is a flow chart which shows the 5th example of this invention, and the configuration of this example is the same as that of the 1st example shown in drawing 2.

[0052] This example when the similarity of a correct answer input voice pattern and the 2nd input voice pattern buffer 62 is larger than the similarity of a correct answer input voice pattern and a dictionary 4. The correct answer description pattern of the 2nd input voice pattern buffer 62 is registered into a dictionary 4. Moreover, when the similarity of a correct answer input voice pattern and the 2nd input voice pattern buffer 62 is smaller, the correct answer description pattern of the 1st and 2nd input voice pattern buffers 61 and 62 is deleted.

[0053] Since the correct answer description pattern currently held at the 1st and 2nd input voice pattern buffers 61 is deleted in this example when the similarity by the side of a dictionary 4 is larger than the similarity by the side of the 2nd input voice pattern buffer 62. When the description pattern of input voice with which input voice was distorted to either of the distortion (deforming), 1st, and 2nd input voice pattern buffers 61 and 62 is held, the description pattern can be deleted.

[0054] In addition, in the above-mentioned example, although similarity is used as a score by collating, distance may be used as a score like the 3rd example. In that case, the sense of the greater than sign shown in drawing 6 turns into reverse sense.

[0055]

[Effect of the Invention] As explained above, when the same voice pattern as the 2nd times is inputted in this invention at the time of renewal of the dictionary by the input voice pattern a recognition result or the input voice pattern which was a correct answer -- or The input voice pattern to which the correct answer

label was given by the user input means, It collates about each with the correct answer standard voice pattern registered into the dictionary 4, and the voice pattern with which the correct answer label currently held temporarily was given to the input voice pattern maintenance means. Since the dictionary 4 is updated when the score between an input voice pattern and the voice pattern held at the input voice pattern maintenance means is more expensive among those collating results While being able to update the standard voice pattern of a dictionary to a new thing whenever it recognizes even if a speaker's utterance condition changes with the passage of time A dictionary can be updated without being restricted to recognition criteria before vocal quality and an utterance situation change, and the high recognition rate corresponding to a speaker's utterance condition can be acquired.

[0056] Moreover, it is avoidable that the voice pattern distorted by chance is registered into a dictionary.

[Translation done.]